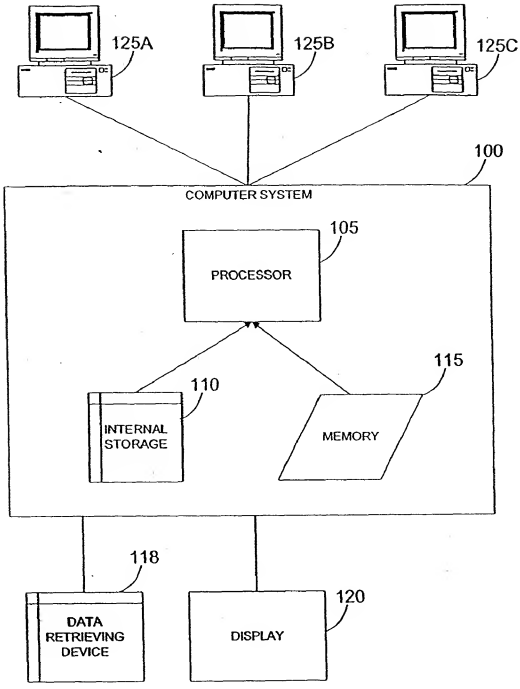
125A
125B
125C
100
COMPUTER SYSTEM
105
PROCESSOR
115
110
INTERNAL STORAGE
MEMORY
118
DATA RETRIEVING DEVICE
120
DISPLAY

FIGURE 1

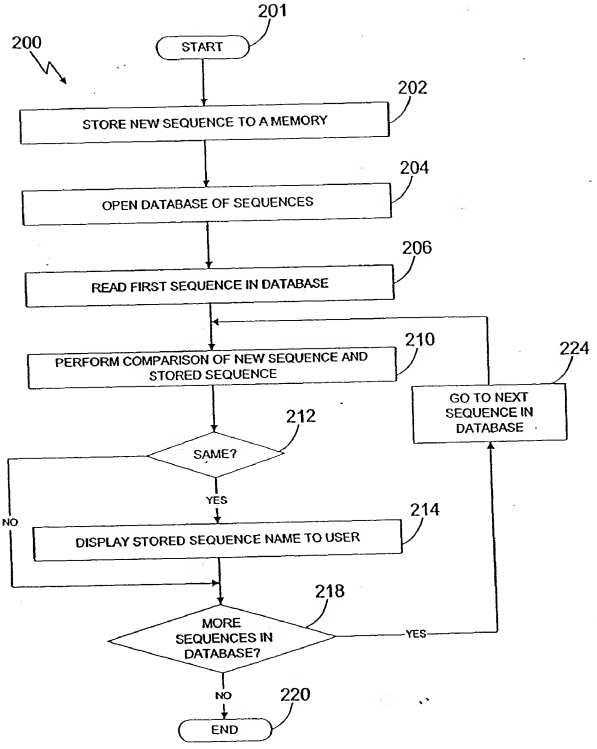
200
201
START
202
STORE NEW SEQUENCE TO A MEMORY
204
OPEN DATABASE OF SEQUENCES
206
READ FIRST SEQUENCE IN DATABASE
210
PERFORM COMPARISON OF NEW SEQUENCE AND STORED SEQUENCE
212
SAME?
YES
NO
214
DISPLAY STORED SEQUENCE NAME TO USER
218
MORE SEQUENCES IN DATABASE?
YES
NO
220
END
224
GO TO NEXT SEQUENCE IN DATABASE

FIGURE 2

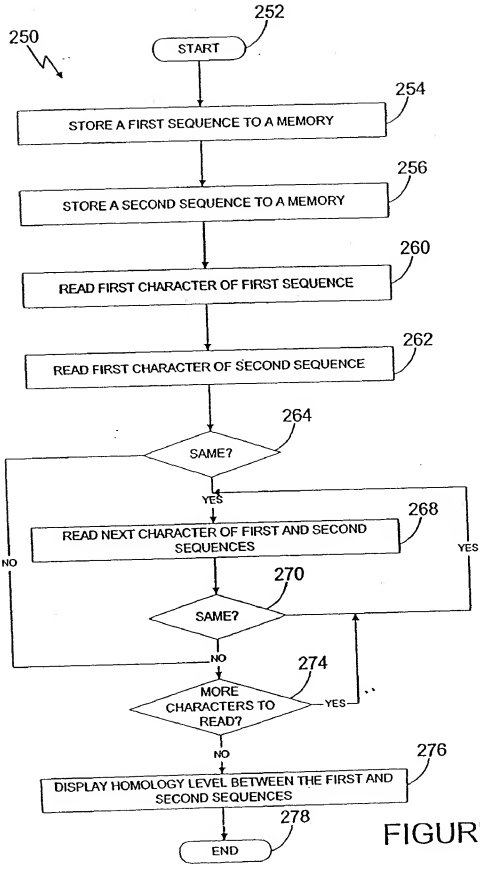
250
252
START
254
STORE A FIRST SEQUENCE TO A MEMORY
256
STORE A SECOND SEQUENCE TO A MEMORY
260
READ FIRST CHARACTER OF FIRST SEQUENCE
262
READ FIRST CHARACTER OF SECOND SEQUENCE
264
SAME?
YES
NO
268
READ NEXT CHARACTER OF FIRST AND SECOND SEQUENCES
YES
270
SAME?
NO
274
MORE CHARACTERS TO READ?
YES
NO
276
DISPLAY HOMOLOGY LEVEL BETWEEN THE FIRST AND SECOND SEQUENCES
278
END

FIGURE 3

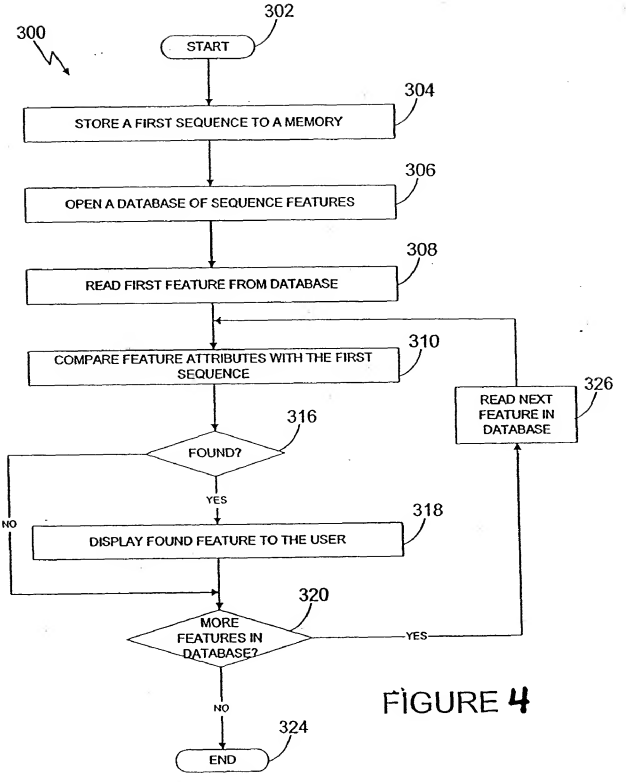
300
302
START
304
STORE A FIRST SEQUENCE TO A MEMORY
306
OPEN A DATABASE OF SEQUENCE FEATURES
308
READ FIRST FEATURE FROM DATABASE
310
COMPARE FEATURE ATTRIBUTES WITH THE FIRST SEQUENCE
316
FOUND?
YES
NO
318
DISPLAY FOUND FEATURE TO THE USER
320
MORE FEATURES IN DATABASE?
YES
NO
324
END
326
READ NEXT FEATURE IN DATABASE

FIGURE 4

FIG. 5

Staphylothermus Marinus - F1-12LC

```
                1 ATG TCT TTA AAC AAG CAC TCT TGG ATG GAT ATG ATA ATA TTT ATT CTC AGC TTT TCT TTC   60
                1 Met Ser Leu Asn Lys His Ser Trp Met Asp Met Ile Ile Phe Ile Leu Ser Phe Ser Phe   20

  6            61 CCA TTA ACA ATG ATC GCA TTA GCT ATC TCT ATG TCG TCA TGG TTT AAT ATA TGG AAT AAT  120
  2            21 Pro Leu Thr Met Ile Ala Leu Ala Ile Ser Met Ser Ser Trp Phe Asn Ile Trp Asn Asn   40

 12           121 GCA TTA AGC GAT CTA GGA CAT GCT GTT AAA AGC AGT GTT GCT CCA ATA TTC AAT CTA GGT  180
  4            41 Ala Leu Ser Asp Leu Gly His Ala Val Lys Ser Ser Val Ala Pro Ile Phe Asn Leu Gly   60

 18           181 CCT GCA ATT GGT GGG ATA CTA ATT GTT ATA GTT GGT TTA AGA AAT CTT TAT TCG TGG AGT  240
  6            61 Leu Ala Ile Gly Gly Ile Leu Ile Val Ile Val Gly Leu Arg Asn Leu Tyr Ser Trp Ser   80

 24           241 AGA GTT AAA GGA TCT TTA ATC ATA TCC ATG GGT GTA TTT CTT AAC TTA ATA GGG GTT TTC  300
  8            81 Arg Val Lys Gly Ser Leu Ile Ile Ser Met Gly Val Phe Leu Asn Leu Ile Gly Val Phe  100

 30           301 GAC GAA GTA TAT GGT TGG ATA CAT TTC CTA GTC TCA GTA TTG TTT TTC TTA TCA ATA ATA  360
 10           101 Asp Glu Val Tyr Gly Trp Ile His Phe Leu Val Ser Val Leu Phe Phe Leu Ser Ile Ile  120

 36           361 GCA TAT TTC ATA GCT ATA TCA ATA CTT GAC AAA TCA TGG ATA GCT GTT CTA CTA ATA ATA  420
 12           121 Ala Tyr Phe Ile Ala Ile Ser Ile Leu Asp Lys Ser Trp Ile Ala Val Leu Leu Ile Ile  140

 42           421 GGT CAT ATT GCA ATG TGG TAT CTA CAC TTT GCT TCA GAG ATT CCG AGA GGT GCG GCT ATT  480
 14           141 Gly His Ile Ala Met Trp Tyr Leu His Phe Ala Ser Glu Ile Pro Arg Gly Ala Ala Ile  160

 48           481 CCC GAG TTA TTA GCG GTA TTC TCG TTT TTA CCA TTC TAT ATA AGA GAC TAT TTT AAA TCA  540
 16           161 Pro Glu Leu Leu Ala Val Phe Ser Phe Leu Pro Phe Tyr Ile Arg Asp Tyr Phe Lys Ser  180

 54           541 TAC ACT AAA CGA TAG                                                            555
 18           181 Tyr Thr Lys Arg End                                                           185
```

FIG. 6A
Pyrodictium - TAG11-17LC

```
1 ATG AAA CTC CTT GAG CCC ACA AAT ACC TCC TAC ACG CTG TTA CAG GAT TTA GCA TTG CAT   60
1 Met Lys Leu Leu Glu Pro Thr Asn Thr Ser Tyr Thr Leu Leu Gln Asp Leu Ala Leu His  20

61 TTT GCA TTT TAC TGG TTT CTG GCC GTG TAT ACG TGG TTA CCC GGT GTC CTA GTC CGG GGC 120
21 Phe Ala Phe Tyr Trp Phe Leu Ala Val Tyr Thr Trp Leu Pro Gly Val Leu Val Arg Gly  40

121 GTA GCT GTG GAC ACA GGG GTG GCT CGG GTG CCT GGG CTC GGC CGG CGC GGT AAG AGG CTG 180
41 Val Ala Val Asp Thr Gly Val Ala Arg Val Pro Gly Leu Gly Arg Arg Gly Lys Arg Leu  60

181 CTC CTG GCC GCT GTG GCT GTC TTG GCG CTT GTT GTG TCC GTT GTT GTC CCG GCT TAT GTG 240
61 Leu Leu Ala Ala Val Ala Val Leu Ala Leu Val Val Ser Val Val Val Pro Ala Tyr Val  80

241 GCG TAT AGT AGT CTG CAC CCG GAG AGC TGT CGG CCC GTT GCG CCG GAG GGG CTC ACC TAC 300
81 Ala Tyr Ser Ser Leu His Pro Glu Ser Cys Arg Pro Val Ala Pro Glu Gly Leu Thr Tyr 100

301 AAA GAG TTC AGC GTG ACC GCG GAG GAT GGC TTG GTG GTT CGG GGC TGG GTG CTG GGC CCC 360
101 Lys Glu Phe Ser Val Thr Ala Glu Asp Gly Leu Val Val Arg Gly Trp Val Leu Gly Pro 120

361 GGC GCT GGG GGC AAC CCG GTG TTC GTT TTG ATG CAC GGG TAT ACT GGG TGC CGC TCG GCG 420
121 Gly Ala Gly Gly Asn Pro Val Phe Val Leu Met His Gly Tyr Thr Gly Cys Arg Ser Ala 140

421 CCC TAC ATG GCT GTG CTG GCC CGG GAG CTC GTG GAG TGG GGG TAC CCG GTG GTT GTG TTC 480
141 Pro Tyr Met Ala Val Leu Ala Arg Glu Leu Val Glu Trp Gly Tyr Pro Val Val Val Phe 160

481 GAC TTC CGG GGC CAC GGG GAG AGC GGG GGC TCG ACG ACG ATT GGG CCC CGG GAG GTG CTG 540
161 Asp Phe Arg Gly His Gly Glu Ser Gly Gly Ser Thr Thr Ile Gly Pro Arg Glu Val Leu 180

541 GAT GCC CGG GCT GTG GTG GGC TAT GTC TCG GAG CGG TTC CCC GGC CGC CGG ATA ATA TTG 600
181 Asp Ala Arg Ala Val Val Gly Tyr Val Ser Glu Arg Phe Pro Gly Arg Arg Ile Ile Leu 200
```

FIG. 6B
Pyrodictium - TAG11-17LC

```
601 GTG GGG TTC AGT ATG GGC GGC GCT GTA GCG ATC GTG GAG GGT GCT GGG GAC CCG CGG GTC  660
201 Val Gly Phe Ser Met Gly Gly Ala Val Ala Ile Val Glu Gly Ala Gly Asp Pro Arg Val  220

661 TAC GCG GTG GCT GCT GAT AGC CCG TAC TAT AGG CTC CGG GAC GTC ATA CCC CGG TGG CTG  720
221 Tyr Ala Val Ala Ala Asp Ser Pro Tyr Tyr Arg Leu Arg Asp Val Ile Pro Arg Trp Leu  240

721 GAG TAC AAG ACG CCG CTG CCG GGC TGG GTG GGT GTG CTG GCC GGG TTC TAC GGG AGG CTG  780
241 Glu Tyr Lys Thr Pro Leu Pro Gly Trp Val Gly Val Leu Ala Gly Phe Tyr Gly Arg Leu  260

781 ATG GCG GGC GTT GAC CTC GGC TTC GGC CCC GCT GGG GTG GAG CGC GTG GAT AAG CCG TTG  840
261 Met Ala Gly Val Asp Leu Gly Phe Gly Pro Ala Gly Val Glu Arg Val Asp Lys Pro Leu  280

841 CTG GTG GTG TAT GGG CCC CGG GAC CCG CTG GTG ACG CGG GAC GAG CGC AGG AGC CTG GCG  900
281 Leu Val Val Tyr Gly Pro Arg Asp Pro Leu Val Thr Arg Asp Glu Ala Arg Ser Leu Ala  300

901 TCC CGT AGC CCG TGT GGC CGT CTC GTC GAG GTT CCT GGG GCT GGC CAC GTG GAG GCC GTG  960
301 Ser Arg Ser Pro Cys Gly Arg Leu Val Glu Val Pro Gly Ala Gly His Val Glu Ala Val  320

961 GAT GTG CTC GGG CCG GGC CGC TAC GCA GAC ATG CTG ATA GAG CTG GCG CAC GAG GAG TGC 1020
321 Asp Val Leu Gly Pro Gly Arg Tyr Ala Asp Met Leu Ile Glu Leu Ala His Glu Glu Cys  340

1021 CCT CCG GGG GCC GGT GGC TGA                                                    1041
 341 Pro Pro Gly Ala Gly Gly End                                                     347
```

FIG. 7A

Archaeoglobus Venificus SN P6-24LC

```
1 ATG CCA TAT GTT AGG AAT GGT GGT GTA AAT ATC TAT TAT GAA CTG GTG GAT GGA CCT GAG   60
1 Met Pro Tyr Val Arg Asn Gly Gly Val Asn Ile Tyr Tyr Glu Leu Val Asp Gly Pro Glu   20

61 CCA CCA ATT GTC TTT GTT CAC GGA TGG ACA GCA AAT ATG AAT TTT TGG AAA GAG CAA AGA  120
21 Pro Pro Ile Val Phe Val His Gly Trp Thr Ala Asn Met Asn Phe Trp Lys Glu Gln Arg   40

121 CGT TAT TTT GCA GGC AGG AAT ATG ATG TTG TTT GTC GAT AAC AGA GGT CAG GGC AGG TCC  180
41 Arg Tyr Phe Ala Gly Arg Asn Met Met Leu Phe Val Asp Asn Arg Gly His Gly Arg Ser   60

181 GAT AAG CCA CTT GGA TAC GAT TTC TAC AGA TTT GAG AAC TTC ATT TCA GAT TTA GAT GCG  240
61 Asp Lys Pro Leu Gly Tyr Asp Phe Tyr Arg Phe Glu Asn Phe Ile Ser Asp Leu Asp Ala   80

241 GTT GTT AGG GAG ACT GGA GTG GAG AAA TTT GTT CTC GTC GGA CAT TCA TTC GGA ACA ATG  300
81 Val Val Arg Glu Thr Gly Val Glu Lys Phe Val Leu Val Gly His Ser Phe Gly Thr Met  100

301 ATC TCT ATG AAG TAC TGT TCG GAG TAT CGG AAT CGG GTT CTT GCT CTA ATC CTC ATA GGT  360
101 Ile Ser Met Lys Tyr Cys Ser Glu Tyr Arg Asn Arg Val Leu Ala Leu Ile Leu Ile Gly  120

361 GGT GGG AGC AGA ATA AAG CTT CTA CAC AGA ATT GGA TAT CCT TTA GCA AAG ATT CTT GCA  420
121 Gly Gly Ser Arg Ile Lys Leu Leu His Arg Ile Gly Tyr Pro Leu Ala Lys Ile Leu Ala  140

421 TCC ATT GCA TAC AAG AAG TCT TCA AGA TTG GTC GCA GAT CTT TCC TTT GGC AAA AAT GCT  480
141 Ser Ile Ala Tyr Lys Lys Ser Ser Arg Leu Val Ala Asp Leu Ser Phe Gly Lys Asn Ala  160

481 GGT GAA CTT AAA GAG TGG GGA TGG AAA CAG GCA ATG GAT TAT ACA CCC TCC TAC GTG GCA  540
161 Gly Glu Leu Lys Glu Trp Gly Trp Lys Gln Ala Met Asp Tyr Thr Pro Ser Tyr Val Ala  160

541 ATG TAC ACG TAC AGA ACT CTA ACG AAA GTG AAT CTT GAA AAT ATC TTG GAG AAA ATA GAC  600
181 Met Tyr Thr Tyr Arg Thr Leu Thr Lys Val Asn Leu Glu Asn Ile Leu Glu Lys Ile Asp  200

601 TGT CCA ACA CTG ATT ATC GTT GGA GAA GAG GAT GCA CTA TTG CCC GTT AGC AAA TCA GTT  660
201 Cys Pro Thr Leu Ile Ile Val Gly Glu Glu Asp Ala Leu Leu Pro Val Ser Lys Ser Val  220
```

FIG.7B
Archaeoglobus Venificus SN P6-24LC

```
661 GAG CTG ACG AGG AGG ATA GAA AAC TCA AAG CTT GTG ATC ATC CCA AAC TCG GGG CAT TGC 720
221 Glu Leu Ser Arg Arg Ile Glu Asn Ser Lys Leu Val Ile Ile Pro Asn Ser Gly His Cys 240

721 GTA ATG CTT GAG AGT CCA AGT GAG GTT AAT AGA GCA ATG GAC GAA TTC ATT TCT TCA GCA 780
241 Val Met Leu Glu Ser Pro Ser Glu Val Asn Arg Ala Met Asp Glu Phe Ile Ser Ser Ala 260

781 CAG TTC TAA                                                                    789
261 Gln Phe End                                                                     263
```

FIG. 8

Aquifax pyrophilus - 28LC

```
   1 TTG AGA TTG AGG AAA TTT GAA GAG ATA AAC CTC GTT CTT TCG GGA GGA GCT GCA AAG GGC   60
   1 Leu Arg Leu Arg Lys Phe Glu Glu Ile Asn Leu Val Leu Ser Gly Gly Ala Ala Lys Gly   20

  61 ATA GCC CAC ATA GGT GTT TTG AAA GCT ATA AAC GAG CTC GGT ATA AGG GTG AGG GCT TTA  120
  21 Ile Ala His Ile Gly Val Leu Lys Ala Ile Asn Glu Leu Gly Ile Arg Val Arg Ala Leu   40

 121 AGC GGG GTG AGC GCC GGG GCA ATC GTT TCG GTC TTT TAT GCC TCA GGC TAC TCC CCT GAA  180
  41 Ser Gly Val Ser Ala Gly Ala Ile Val Ser Val Phe Tyr Ala Ser Gly Tyr Ser Pro Glu   60

 181 GGG ATG TTC AGC CTT CTG AAG AGG GTA AAC TGG CTG AAG CTG TTT AAG TTC AAG CCA CCT  240
  61 Gly Met Phe Ser Leu Leu Lys Arg Val Asn Trp Leu Lys Leu Phe Lys Phe Lys Pro Pro   80

 241 CTG AAG GGA TTG ATA GGG TGG GAG AAG GCT ATA AGA TTC CTT GAG GAA GTT CTC CCT TAC  300
  81 Leu Lys Gly Leu Ile Gly Trp Glu Lys Ala Ile Arg Phe Leu Glu Glu Val Leu Pro Tyr  100

 301 AGG AGA ATA GAA AAA CTT GAG ATA CCG ACG TAT ATA TGC GCG ACG GAT TTA TAC TCG GGA  360
 101 Arg Arg Ile Glu Lys Leu Glu Ile Pro Thr Tyr Ile Cys Ala Thr Asp Leu Tyr Ser Gly  120

 361 AGG GCT CTA TAC CTC TCG GAA GGG AGT TTA ATC CCC GCA CTT CTC GGC AGC TGT GCA ATT  420
 121 Arg Ala Leu Tyr Leu Ser Glu Gly Ser Leu Ile Pro Ala Leu Leu Gly Ser Cys Ala Ile  140

 421 CCC GGC ATA TTT GAA CCC GTT GAG TAT AAG AAT TAC TTG CTC GTT GAC GGA GGT ATA GTT  480
 141 Pro Gly Ile Phe Glu Pro Val Gly Tyr Lys Asn Tyr Leu Leu Val Asp Gly Gly Ile Val  160

 481 AAC AAC CTT CCC GTT GAG CCC TTT CAG GAA AGC GGT ATT CCC ACC GTT TGC GTT GAT GTC  540
 161 Asn Asn Leu Pro Val Glu Pro Phe Gln Glu Ser Gly Ile Pro Thr Val Cys Val Asp Val  180

 541 CTT CCC ATA GAG CCG GAA AAG GAT ATA AAG AAC ATT CTT CAC ATC CTT TTG AGG AGC TTC  600
 181 Leu Pro Ile Glu Pro Glu Lys Asp Ile Lys Asn Ile Leu His Ile Leu Leu Arg Ser Phe  200

 601 TTT CTT GCG GTC CGC TCA AAC TCC GAA AAG AGA AAG GAG TTT TGT GAC CTC GTT ATA GTT  660
 201 Phe Leu Ala Val Arg Ser Asn Ser Glu Lys Arg Lys Glu Phe Cys Asp Leu Val Ile Val  220

 661 CCT GAG CTT GAG GAG TTC ACA CCC CTT GAT GTT AGA AAA GCG GAC CAA ATA ATG GAG AGG  720
 221 Pro Glu Leu Glu Glu Phe Thr Pro Leu Asp Val Arg Lys Ala Asp Gln Ile Met Glu Arg  240

 721 GGA TAC ATA AAG GCC TTA GAG GTA CTT TCT GAA TAG                                  756
 241 Gly Tyr Ile Lys Ala Leu Glu Val Leu Ser Glu End                                  252
```

FIG. 9A
M11TL-29L.

```
  1 ATG TTT AAT ATC AAT GTC TTT GTT AAT ATA TCT TGG CTG TAT TTT TCA GGG ATA GTT ATG   60
  1 Met Phe Asn Ile Asn Val Phe Val Asn Ile Ser Trp Leu Tyr Phe Ser Gly Ile Val Met  20

 61 AAG ACT GTG GAA GAG TAT GCG CTA CTT GAA ACA GGC GTA AGA GTG TTT TAT CGG TGT GTA  120
 21 Lys Thr Val Glu Glu Tyr Ala Leu Leu Glu Thr Gly Val Arg Val Phe Tyr Arg Cys Val  40

121 ATC CCG GAG AAA GCT TTT AAC ACT TTG ATA ATA GGT TCA CAC GGA TTG GGG GCG CAC AGT  180
 41 Ile Pro Glu Lys Ala Phe Asn Thr Leu Ile Ile Gly Ser His Gly Leu Gly Ala His Ser  60

181 GGA ATC TAC ATT AGT GTT GCT GAA GAA TTT GCT AGG CAC GGA TTT GGA TTC TGC ATG CAC  240
 61 Gly Ile Tyr Ile Ser Val Ala Glu Glu Phe Ala Arg His Gly Phe Gly Phe Cys Met His  80

241 GAT CAA AGG GGA CAT GGG AGA ACG GCA AGC GAT AGA GAA AGA GGG TAT GTG GAG GGC TTT  300
 81 Asp Gln Arg Gly His Gly Arg Thr Ala Ser Asp Arg Glu Arg Gly Tyr Val Glu Gly Phe 100

301 CAC AAC TTC ATA GAG GAT ATG AAG GCC TTC TCC GAT TAT GCC AAG TGG CGC GTG GGA GGT  360
101 His Asn Phe Ile Glu Asp Met Lys Ala Phe Ser Asp Tyr Ala Lys Trp Arg Val Gly Gly 120

361 GAC GAA ATA ATA TTG CTA GGA CAC AGT ATG GGC GGG CTG ATA GCG CTC TTA ACA GTT GCA  420
121 Asp Glu Ile Ile Leu Leu Gly His Ser Met Gly Gly Leu Ile Ala Leu Leu Thr Val Ala 140

421 ACT TAT AAA GAA ATC GCC AAG GGA GTT ATC GCG CTA GCC CCG GCC CTC CAA ATC CCC TTA  480
141 Thr Tyr Lys Glu Ile Ala Lys Gly Val Ile Ala Leu Ala Pro Ala Leu Gln Ile Pro Leu 160

481 ACC CCG GCT AGA AGA CTT GTT CTA AGC CTC GCG TCA AGG CTT GCC CCG CAT TCT AAG ATC  540
161 Thr Leu Gln Arg Arg Leu Val Leu Ser Leu Ala Ser Arg Leu Ala Pro His Ser Lys Ile 180

541 ACC TTA CAA AGG AGA TTG CCG CAG AAA CCA GAG GGT TTT CAA AGA GCA AAA GAT ATA GAA  600
181 Thr Leu Gln Arg Arg Leu Pro Gly Lys Pro Glu Gly Phe Gln Arg Ala Lys Asp Ile Glu 200

601 TAC AGT CTG AGT GAA ATA TCA GTC AAG CTC GTG GAC GAA ATG ATT AAA GCA TCA TCT ATG  660
201 Tyr Ser Leu Ser Glu Ile Ser Val Lys Leu Val Asp Glu Met Ile Lys Ala Ser Ser Met 220

661 TTG TGG ACC ATA GCA GGG GAA ATT AAT ACT CCC GTC CTG CTT ATT CAT GGG GAA AAA GAC 720
221 Phe Trp Thr Ile Ala Gly Glu Ile Asn Thr Pro Val Leu Leu Ile Lys Ala Ser Ser Met 240
```

FIG. 9 B
M11TL-29L.

```
721 AAT GTC ATA CCT CCG GAG GCG AGC AAA AAA GCC TAC CAA TTA ATA CCT TCA TTC CCT AAA  780
241 Asn Val Ile Pro Pro Glu Ala Ser Lys Lys Ala Tyr Gln Leu Ile Pro Ser Phe Pro Lys

781 GAG TTG AAA ATA TAC CCC GAT CTT GGA CAC AAC TTG TTT TTT GAA CCA GGC GCG GTG AAA  840
261 Glu Leu Lys Ile Tyr Pro Asp Leu Gly His Asn Leu Phe Phe Glu Pro Gly Ala Val Lys  280

841 ATC GTC ACA GAC ATT GTA GAG TGG GTT AAG AAT CTA CCC AGG GAA AAT CCT TAA  894
281 Ile Val Thr Asp Ile Val Glu Trp Val Lys Asn Leu Pro Arg Glu Asn Pro End  298
```

FIG. 10A
***Thermococcus* CL-2-30LC**

```
  1 ATG GAG GTT TAC AAG GCC AAA TTC GGC GAA GCA AAG CTC GGC TGG GTC GTT CTG GTT CAT  60
  1 Met Glu Val Tyr Lys Ala Lys Phe Gly Glu Ala Lys Leu Gly Trp Val Val Leu Val His  20

 61 GGC CTC GGC GAG CAC AGC GGA AGG TAT GGA AGA CTG ATT AAG GAA CTC AAC TAT GCC GGC 120
 21 Gly Leu Gly Glu His Ser Gly Arg Tyr Gly Arg Leu Ile Lys Glu Leu Asn Tyr Ala Gly  40

121 TTT GGA GTT TAC ACC TTC GAC TGG CCC GGC CAC GGG AAG AGC CCG GGC AAG AGA GGG CAC 180
 41 Phe Gly Val Tyr Thr Phe Asp Trp Pro Gly His Gly Lys Ser Pro Gly Lys Arg Gly His  60

181 ACG AGC GTC GAG GAG GCG ATG GAA ATC ATC GAC TCG ATA ATC GAG GAG ATC AGG GAG AAG 240
 61 Thr Ser Val Glu Glu Ala Met Glu Ile Ile Asp Ser Ile Ile Glu Glu Ile Arg Glu Lys  80

241 CCC TTC CTC TTC GGC CAC AGC CTC GGT GGT CTA ACT GTC ATC AGG TAC GCT GAG ACG CGG 300
 81 Pro Phe Leu Phe Gly His Ser Leu Gly Gly Leu Thr Val Ile Arg Tyr Ala Glu Thr Arg 100

301 CCC GAT AAA ATA CGG GGA TTA ATA GCT TCC TCG CCT GCC CTC GCC AAG AGC CCG GAA ACG 360
101 Pro Asp Lys Ile Arg Gly Leu Ile Ala Ser Ser Pro Ala Leu Ala Lys Ser Pro Glu Thr 120

361 CCG GGC TTC ATG GTG GCC CTC GCG AAG TTC CTT GGA AAG ATC GCC CCG GGA GTT GTT CTC 420
121 Pro Gly Phe Met Val Ala Leu Ala Lys Phe Leu Gly Lys Ile Ala Pro Gly Val Val Leu 140

421 TCC AAC GGC ATA AAG CCG GAA CTC CTC TCG AGG AAC AGG GAC GCC GTG AGG AGG TAC GTT 480
141 Ser Asn Gly Ile Lys Pro Glu Leu Leu Ser Arg Asn Arg Asp Ala Val Arg Arg Tyr Val 160

481 GAA GAC CCA CTC GTC CAC GAC AGG ATT TCG GCC AAG CTG GGA AGG AGC ATC TTC GTG AAC 540
161 Glu Asp Pro Leu Val His Asp Arg Ile Ser Ala Lys Leu Gly Arg Ser Ile Phe Val Asn 180

541 ATG GAG CTG GCC CAC AGG GAG GCG GAC AAG ATA AAA GTC CCG ATC CTC CTT CTG ATG GGC 600
181 Met Glu Leu Ala His Arg Glu Ala Asp Lys Ile Lys Val Pro Ile Leu Leu Leu Ile Gly 200

601 ACT GGC GAT GTA ATA ACC CCG CCT GAA GGC TCA CGC AGA CTC TTC GAG GAG CTG GCC GTC 660
201 Thr Gly Asp Val Ile Thr Pro Pro Glu Gly Ser Arg Arg Leu Phe Glu Glu Leu Ala Val 220

661 GAG AAC AAA ACC CTG AGG GAG TTC GAG GGG GCG TAC CAC GAG ATA TTT GAA GAC CCC GAG 720
221 Glu Asn Lys Thr Leu Arg Glu Phe Glu Gly Ala Tyr His Glu Ile Phe Glu Asp Pro Glu 240
```

FIG.10B

Thermococcus CL-2-30LC

```
721 TGG GCC GAG GAG TTC CAC GAA ACA ATT GTT AAG TGG CTG GTT GAA AAA TCG TAC TCT TCG  780
241 Trp Ala Glu Glu Phe His Glu Thr Ile Val Lys Trp Leu Val Glu Lys Ser Tyr Ser Ser  260

781 GCT CAA TAA  789
261 Ala Gln End  263
```

FIG. 11
Aquifex VF5-34LC

```
   1 TTG ATT GGC AAT TTG AAA TTG AAG AGG TTT GAA GAG GTT AAC TTA GTT CTT TCG GGA GGG   60
   1 Leu Ile Gly Asn Leu Lys Leu Lys Arg Phe Glu Glu Val Asn Leu Val Leu Ser Gly Gly   20

  61 GCT GCC AAG GGT ATC GCC CAT ATA GGT GTT TTA AAA GCT CTG GAA GAG CTC GGT ATA AAG  120
  21 Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys Ala Leu Glu Glu Leu Gly Ile Lys   40

 121 GTA AAG AGG CTC AGC GGG GTA AGT GCT GGA GCT ATC GTT TCC GTC TTT TAC GCT TCG GGC  180
  41 Val Lys Arg Leu Ser Gly Val Ser Ala Gly Ala Ile Val Ser Val Phe Tyr Ala Ser Gly   60

 181 TAC ACT CCC GAC GAG ATG TTA AAA CTC CTG AAA GAG GTA AAC TGG CTC AAA CTT TTT AAG  240
  61 Tyr Thr Pro Asp Glu Met Leu Lys Leu Leu Lys Glu Val Asn Trp Leu Lys Leu Phe Lys   80

 241 TTC AAA ACA CCG AAA ATG GGC TTA ATG GGG TGG GAG AAG GCT GCA GAG TTT TTG GAA AAA  300
  81 Phe Lys Thr Pro Lys Met Gly Leu Met Gly Trp Glu Lys Ala Ala Glu Phe Leu Glu Lys  100

 301 GAG CTC GGA GTT AAG AGG CTG GAA GAC CTG AAC ATA CCA ACC TAT CTT TGC TCG GCG GAT  360
 101 Glu Leu Gly Val Lys Arg Leu Glu Asp Leu Asn Ile Pro Thr Tyr Leu Cys Ser Ala Asp  120

 361 CTG TAC ACG GGA AAG GCT CTT TAC TTC GGC AGA GGT GAC TTA ATT CCC GTG CTT CTC GGA  420
 121 Leu Tyr Thr Gly Lys Ala Leu Tyr Phe Gly Arg Gly Asp Leu Ile Pro Val Leu Leu Gly  140

 421 AGT TGT TCC ATA CCC GGG ATT TTT GAA CCA GTT GAG TAC GAG AAT TTT CTA CTT GTT GAC  480
 141 Ser Cys Ser Ile Pro Gly Ile Phe Glu Pro Val Glu Tyr Glu Asn Phe Leu Leu Val Asp  160

 481 GGA GGT ATA GTG AAC AAC CTG CCC GTA GAA CCT TTG GAA AAG TTC AAA GAA CCC ATA ATC  540
 161 Gly Gly Ile Val Asn Asn Leu Pro Val Glu Pro Leu Glu Lys Phe Lys Glu Pro Ile Ile  180

 541 GGG GTA GAT GTG CTT CCC ATA ACT CAA GAA AGA AAG ATT AAA AAT ATA CTC CAC ATC CTT  600
 181 Gly Val Asp Val Leu Pro Ile Thr Gln Glu Arg Lys Ile Lys Asn Ile Leu His Ile Leu  200

 601 ATA AGG AGC TTC TTT CTG GCG GTT CGT TCC AAT TCG GAA AAG AGA AAG GAG TTC TGC AAC  660
 201 Ile Arg Ser Phe Phe Leu Ala Val Arg Ser Asn Ser Glu Lys Arg Lys Glu Phe Cys Asn  220

 661 GTA GTT ATA GAA CCT CCC CTT GAA GAG TTC TCT CCT CTG GAC GTA AAT AAG GCG GAC GAG  720
 221 Val Val Ile Glu Pro Pro Leu Glu Glu Phe Ser Pro Leu Asp Val Asn Lys Ala Asp Glu  240

 721 ATA TTC TGC GGG GAT ATG AGA GCA CTT TAA  750
 241 Ile Phe Cys Gly Asp Met Arg Ala Leu End  250
```

FIG.12A

Teredinibacter - 42L

```
1 ATG CCA GCT AAT GAC TCA CCC ACG ATC GAC TTT AAT CCT CGC GGC ATT CTT CGC AAC GCT   60
1 Met Pro Ala Asn Asp Ser Pro Thr Ile Asp Phe Asn Pro Arg Gly Ile Leu Arg Asn Ala  20

61 CAC GCA CAG GTT ATT TTA GCG ACT TCC GGC TTG CGC AAA GCG TTT TTG AAA CGC ACG CAC 120
21 His Ala Gln Val Ile Leu Ala Thr Ser Gly Leu Arg Lys Ala Phe Leu Lys Arg Thr His  40

121 AAG AGC TAC CTC AGC ACT GCC CAA TGG CTG GAG CTC GAT GCC GGC AAC GGA GTT ACC TTG 180
41 Lys Ser Tyr Leu Ser Thr Ala Gln Trp Leu Glu Leu Asp Ala Gly Asn Gly Val Thr Leu  60

181 GCC GGA GAG CTT AAC ACA GCG CCT GCA ACT GCA TCC TCC TCC CAC CCG GCG CAC AAG AAC 240
61 Ala Gly Glu Leu Asn Thr Ala Pro Ala Thr Ala Ser Ser Ser His Pro Ala His Lys Asn  80

241 ACT CTG GTT ATT GTG CTG CAC GGC TGG GAA GGC TCC AGC CAG TCG GCC TAT GCG ACC TCC 300
81 Thr Leu Val Ile Val Leu His Gly Trp Glu Gly Ser Ser Gln Ser Ala Tyr Ala Thr Ser 100

301 GCT GGC AGC ACG CTT TTC GAC AAT GGG TTC GAC ACT TTT CGC CTT AAT TTT CGC GAT CAC 360
101 Ala Gly Ser Thr Leu Phe Asp Asn Gly Phe Asp Thr Phe Arg Leu Asn Phe Arg Asp His 120

361 GGC GAC ACC TAC CAC TTA AAC CGC GGC ATA TTT AAC TCA TCG CTG ATT GAC GAA GTA GTG 420
121 Gly Asp Thr Tyr His Leu Asn Arg Gly Ile Phe Asn Ser Ser Leu Ile Asp Glu Val Val 140

421 GGC GCA GTC AAA GCC ATC CAG CAG CAA ACC GAC TAC GAC AAG TAT TGC CTG ATG GGG TTC 480
141 Gly Ala Val Lys Ala Ile Gln Gln Gln Thr Asp Tyr Asp Lys Tyr Cys Leu Met Gly Phe 160

481 TCA CTG GGT GGG AAC TTT GCC TTG CGC GTC GCG GTG CGG GAA CAG CAT CTC GCT AAA CCG 540
161 Ser Leu Gly Gly Asn Phe Ala Leu Arg Val Ala Val Arg Glu Gln His Leu Ala Lys Pro 180

541 CTA GCG GGC GTG CTC GCC GTA TGC CCG GTA CTC GAC CCC GCA CAC ACC ATG ATG GCC CTA 600
181 Leu Ala Gly Val Leu Ala Val Cys Pro Val Leu Asp Pro Ala His Thr Met Met Ala Leu 200

601 AAC CGA GGT GCG TTT TTC TAC GGC CGC TAT TTT GCG CAT AAA TGG AAG CGC TCG TTAACC 660
201 Asn Arg Gly Ala Phe Phe Tyr Gly Arg Tyr Phe Ala His Lys Trp Lys Arg Ser Leu Thr 220

661 GCA AAA CTT GCA GCT TTC CCA GAC TAC AAA TAC GGC AAA GAT TTA AAA TCG ATA CAC ACG 720
221 Ala Lys Leu Ala Ala Phe Pro Asp Tyr Lys Tyr Gly Lys Asp Leu Lys Ser Ile His Thr 240
```

FIG. 12B
Teredinibacter - 42L

```
721 CTT GAT GAG TTA AAC AAC TAT TTC ATT CCC CGC TAC ACC GGC TTC AAC TCA GTC TCC GAA  780
241 Leu Asp Glu Leu Asn Asn Tyr Phe Ile Pro Arg Tyr Thr Gly Phe Asn Ser Val Ser Glu  260

781 TAC TTC AAA AGT TAC ACG CTC ACC GGG CAG AAG CTC GCG TTT CTC AAC TGC CCC AGT TAC  840
261 Tyr Phe Lys Ser Tyr Thr Leu Thr Gly Gln Lys Leu Ala Phe Leu Asn Cys Pro Ser Tyr  280

841 ATT CTG GCA GCT GGC GAC GAC CCA ATA ATT CCA GCA TCC GAC TTT CAG AAA ATA GCC AAG  900
281 Ile Leu Ala Ala Gly Asp Asp Pro Ile Ile Pro Ala Ser Asp Phe Gln Lys Ile Ala Lys  300

901 CCT GCG AAT CTG CAC ATA ACA GTA ACG CAA CAA GGT TCT CAT TGC GCA TAC CTG GAA AAC  960
301 Pro Ala Asn Leu His Ile Thr Val Thr Gln Gln Gly Ser His Cys Ala Tyr Leu Glu Asn  320

961 CTG CAT AAA CCT AGT GCT GCC GAC AAA TAT GCG GTG AAA TTA TTT GGA GCC TGT TGA 1017
321 Leu His Lys Pro Ser Ala Ala Asp Lys Tyr Ala Val Lys Leu Phe Gly Ala Cys End  339
```

FIG.13A

Archeoglogus fulgidas VC16 - 16MC1

```
ATG CTT GAT ATG CCA ATC GAC CCT GTT TAC TAC CAG CTT GCT GAG TAT
Met Leu Asp Met Pro Ile Asp Pro Val Tyr Tyr Gln Leu Ala Glu Tyr
 1               5                  10                  15

TTC GAC AGT CTG CCG AAG TTC GAC CAG TTT TCC TCG GCC AGA GAG TAC
Phe Asp Ser Leu Pro Lys Phe Asp Gln Phe Ser Ser Ala Arg Glu Tyr
            20                  25                  30

AGG GAG GCG ATA AAT CGA ATA TAC GAG GAG AGA AAC CGG CAG CTG AGC
Arg Glu Ala Ile Asn Arg Ile Tyr Glu Glu Arg Asn Arg Gln Leu Ser
        35                  40                  45

CAG CAT GAG AGG GTT GAA AGA GTT GAG GAC AGG ACG ATT AAG GGG AGG
Gln His Glu Arg Val Glu Arg Val Glu Asp Arg Thr Ile Lys Gly Arg
    50                  55                  60

AAC GGA GAC ATC AGA GTC AGA GTT TAC CAG CAG AAG CCC GAT TCC CCG
Asn Gly Asp Ile Arg Val Arg Val Tyr Gln Gln Lys Pro Asp Ser Pro
65                  70                  75                  80

GGT CTG GTT TAC TAT CAC GGT GGT GGA TTT GTG ATT TGC AGC ATC GAG
Val Leu Val Tyr Tyr His Gly Gly Gly Phe Val Ile Cys Ser Ile Glu
                85                  90                  95

TCG CAC GAC GCC TTA TGC AGG AGA AYY GCG AGA CTT TCA AAC TCT ACC
Ser His Asp Ala Leu Cys Arg ARg Ile Ala Arg Leu Ser Asn Ser Thr
            100                 105                 110

GTA GTC TCC GTG GAT TAC AGG CTC GCT CCT GAG CAC AAG TTT CCC CCC
Val Val Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Ala
        115                 120                 125

CCA GTT TAT CAT TGC TAC GAT GCG ACC AAG TGG GTT GCT GAG AAC CGG
Ala Val Tyr Asp Cys Tyr Aso Ala Thr Lys Trp Val Ala Glu Asn Ala
    130                 135                 140

GAG GAG CTG AGG ATT GAC CCG TCA AAA ATC TTC GTT GGG GGG GAC AGT
Glu Glu Leu Arg Ile Asp Pro Ser Lys Ile Phe Val Gly Gly Asp Ser
145                 150                 155                 160

GCG GGA CGG AAT CTT GCC CCG GCG CTT TCA ATA ATG GCG AGA GAC AGC
Ala Gly Gly Asn Leu Ala Ala Ala Val Ser Ile Met Ala Arg Asp Ser
                165                 170                 175
```

FIG. 13B
Archeoglogus fulgidas VC16 - 16MC1

```
GGA GAA GAT TTC ATA AAG CAT GAA ATT CTA ACT TAC CCC GTT GTG AAC
Gly Glu Asp Phe Ile Lys His Gln Ile Leu Ile Tyr Pro Val Val Asn
          180                 185                 190
TTT GTA GCC CCC ACA CCA TCG CTT CTG GAG TTT GGA GAG GGG CTG TGG
Phe Val Ala Pro Thr Pro Ser Leu Leu Glu Phe GLy Glu Gly Leu Trp
     195                 200                 205
ATT CTC GAC CAG AAG ATA ATG AGT TGG TTC TCG GAG CAG TAC TTC TCC
Ile Leu Asp Gln Lys Ile Met Ser Trp Phe Ser Glu Gln Tyr Phe Ser
210                 215                 230
AGA GAG GAA GAT AAG TTC AAG CCC CTC GCC TCC GTA ATC TTT GCG GAC
Arg Glu Glu Aso Lys Phe Asn Pro Leu Ala Ser Val Ile Phe Ala Asp
235                 240                 245                 250
CTT GAG AAC CTA CCT CCT GTG CTG ATC ATA ACC GCC GAA TAC GAC CCG
Leu Glu Asn Leu Pro Pro Ala Leu Ile Ile Thr Ala Glu Tyr Asp Pro
          255                 260                 265
CTG AGA GAT GAA GGA GAA GTT TTC GGG CAG ATG CTG AGA AGA GCC GGT
Leu Arg Asp Glu Gly Glu Val Phe Gly Gln Met Leu Arg Arg Ala Gly
     270                 275                 280
GTT GAG GCG AGC ATC GTC AGA TAC AGA GGC GTG CTT CAC GGA TTC ATC
Val Glu Ala Ser Ile Val Arg Tyr Arg Gly Val Leu His Gly Phe Ile
285                 290                 295
AAT TAC TAT CCC GTG CTG AAG GCT GCG AGG GAT GCG ATA AAC CAG ATT
Asn Tyr Tyr Pro Val Leu Lys Ala Ala Arg Asp Ala Ile Asn Gln Ile
300                 305                 310
GCC GCT CTT CTT GTG TTC GAC TAG
Ala Ala Leu Ieu Val Phe Asp
315                 320
```

FIG. 14A
Sulfolobus Solfatarious P1 8LC1

```
 ATG CCC CTA GAT CCT AGA ATT AAA AAG TTA CTA GAA TCA GCT CTT ACT
Met Pro Leu Asp Pro Arg Ile Lys Lys Leu Leu Glu Ser Ala Leu Thr
  1               5                  10                  15
ATA CCA ATT GGT AAA GCC CCA GTA GAA GAG GTA AGA AAG ATA TTT AGG
Ile Pro Ile Gly Lys Ala Pro Val Glu Glu Val Arg Lys Ile Phe Arg
            20                  25                  30
CAA TTA GCG TCG GCA GCT CCC AAA GTC GAA GTT GGA AAA GTA GAA GAT
Gln Leu Ala Ser Ala Ala Pro Lys Val Glu Val Gly Lys Val Glu Asp
        35                  40                  45
ATA AAA ATA CCA GGC AGT GAA ACC GTT ATA AAC GCT AGA GTG TAT TTT
Ile Lys Ile Pro Gly Ser Glu Thr Val Ile Asn Ala Arg Val Tyr Phe
    50                  55                  60
CCG AAG AGT AGC GGT CCT TAT GGT GTT CTA GTG TAT CTT CAT GGA GGC
Pro Lys Ser Ser Gly Pro Tyr Gly Val Leu Val Tyr Leu His Gly Gly
65                  70                  75                  80
GGT TTT GTA ATA GGC GAT GTG GAA TCT TAT GAC CCA TTA TGT AGA GCA
Gly Phe Val Ile Gly Asp Val Glu Ser Tyr Asp Pro Leu Cys Arg Ala
                85                  90                  95
ATT ACA AAT GCG TGC AAT TGC GTT GTA GTA TCA GTG GAC TAT AGG TTA
Ile Thr Asn Ala Cys Asn Cys Val Val Val Ser Val Asp Tyr Arg Leu
            100                 105                 110
GCT CCA GAA TAC AAG TTT CCT TCT GCA GTT ATC GAT TCA TTT GAC GCT
Ala Pro Glu Tyr Lys Phe Pro Ser Ala Val Ile Asp Ser Phe Asp Ala
        115                 120                 125
ACT AAT TGG GTT TAT AAC AAT TTA GAT AAA TTT GAT GGA AAG ATG GGA
Thr Asn Trp Val Tyr Asn Asn Leu Asp Lys Phe Asp Gly Lys Met Gly
    130                 135                 140
GTT GCG ATT GCG GGA GAT AGT GCT GGA GGA AAT TTG GCA GCG GTT GTA
Val Ala Ile Ala Gly Asp Ser Ala Gly Gly Asn Leu Ala Ala Val Val
145                 150                 155                 160
```

FIG. 14B

Sulfolobus Solfatarious P1 8LC1

```
 GCT CTT CTT TCA AAG GGT AAA ATT AAT TTG AAG TAT CAA ATA CTG GTT
Ala Leu Leu Ser Lys Gly Lys Ile Asn Leu Lys Tyr Gln Ile Leu Val
                165                 170                 175
 TAC CCA GCG GTA AGT TTA GAT AAC GTT TCA AGA TCC ATG ATA GAG TAC
Tyr Pro Ala Val Ser Leu Asp Asn Val Ser Arg Ser Met Ile Glu Tyr
            180                 185                 190
TCT GAT GGG TTC TTC CTT ACC AGA GAG CAT ATA GAG TGG TTC GGT TCT
Ser Asp Gly Phe Phe Leu Thr Arg Glu His Ile Glu Trp Phe Gly Ser
        195                 200                 205
CAA TAC TTA CGA AGC CCT GCA GAT TTG CTA GAC TTT AGG TTC TCT CCA
Gln Tyr Leu Arg Ser Pro Ala Asp Leu Leu Asp Phe Arg Phe Ser Pro
    210                 215                 220
ATT CTG GCG CAA GAT TTC AAC GGA TTA CCT CCA GCC TTC ATA ATA ACA
Ile Leu Ala Gln Asp Phe Asn Gly Leu Pro Pro Ala Leu Ile Ile Thr
225                 230                 235                 240
GCA GAA TAC GAT CCA GTA AGG GAT CAA GGA GAA GCG TAT GCA AAT AAA
Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Ala Tyr Ala Asn Lys
                245                 250                 255
CTA CTA CAA GCT GGA GTC TCA GTT ACT AGT GTG AGA TTT AAC AAC GTT
Leu Leu Gln Ala Gly Val Ser Val Thr Ser Val Arg Phe Asn Asn Val
            260                 265                 270
ATA CAC GGA TTC CTC TCA TTC TTT CCG TTG ATG GAG CAA GGA AGA GAT
Ile His Gly Phe Leu Ser Phe Phe Pro Leu Met Glu Gln Gly Arg Asp
        275                 280                 285
GCT ATA GGT CTG ATA GGG TCT GTG TTA AGA CGA GTA TTT TAT GAT AAA
Ala Ile Gly Lru Ile Gly Ser Val Leu Arg Arg Val Phe Tyr Asp Lys
    290                 295                 300
ATT TAA
Ile
305
```